

INDEPENDENT COMPONENTS OF ODOUR SIGNALS

*Martin Kermit**

Agricultural University of Norway
Physics Section
N-1432 Ås, Norway

Oliver Tomic†

MATFORSK
Norwegian food research institute
N-1432 Ås, Norway

ABSTRACT

If two independent observations or processes are measured with the same apparatus, the inherent nature of the measuring device will in many cases introduce a dependency between the two recorded processes object to inspection. In this paper a suggestion of how Independent Component Analysis (ICA) can be used to identify such device dependencies and in turn give an estimated reconstruction of the observations without the correlation between signals introduced by the apparatus. The procedure is illustrated with the use of an "electronic nose" used to sample odours from mixtures of alcohol solutions. It is shown that ICA as a novel tool in the analysis of odour signals can extract the independent odour sources and give acceptable estimates of the ratio with which the alcohol solutions were mixed with two different approaches.

1. INTRODUCTION

Independent Component Analysis (ICA) has successfully been applied as a solution to the so called cocktail party problem and blind source separation problems in general. A necessary assumption for robust estimation of source signals is that the sensors used to record the signals do not introduce any new, significant dependencies between the sources that were not present before the signal reached the recording device. By the use of sensors with varying characteristics, a satisfactory reconstruction of the source signals can not be guaranteed by the conventional use of ICA. A general sampling system will in principle never be able to exactly record a random process with no influence introduced by the sensors or the nature of the measuring device used in the recording process.

If it is assumed that sensors dependencies can be modeled as mixtures governed by the nature of the recording process, it will in this paper be shown that estimates of the source signals, sensor dependencies and the unmixing system is possible.

The next section presents the two ICA-models that are used in this work. Section 3 illuminates the problem with correlated signals by the use of a generated set of signals where the correlation is known. Thereafter, in Section 4, the proposed methods are applied on a real world example with signals sampled by an electronic nose. In the last section, a conclusion and some further guidelines are given.

2. ICA-MODELS

Several introductions to ICA exist. Among others, the references [1, 2] are recommended for the reader not familiar to ICA. Here, only a brief review of the conventional ICA-model, followed by a proposal of an extended ICA-model, will be given.

2.1. ICA-model of independent sources

The following ICA model is based on the work reported by Comon [3] and Hyvärinen & Oja [4].

Suppose that N observations $x_1, x_2, \dots, x_n, \dots, x_N$ of M independent sources $s_1, s_2, \dots, s_m, \dots, s_M$ have been linearly mixed together as

$$x_n = a_{n1}s_1 + \dots + a_{nm}s_m + \dots + a_{nM}s_M, \quad (1)$$

for $n = 1, 2, \dots, N$. The mixing system is then characterized by the set of coefficients a_{nm} , under the assumption that both the mixtures x_n and the sources s_m are random variables. It is convenient to use matrix notation for $\mathbf{x} = [x_1, x_2, \dots, x_N]^T$ and $\mathbf{s} = [s_1, s_2, \dots, s_M]^T$, so that the mixing system in Eq. (1) can be rewritten

$$\mathbf{x} = \mathbf{A}\mathbf{s}, \quad (2)$$

where \mathbf{A} is a full rank $N \times M$ scalar matrix containing the coefficients a_{nm} of the mixing system. This statistical model is called an ICA model [5], and the goal is to estimate an inverse of \mathbf{A} giving a system \mathbf{W} such that

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \quad (3)$$

thus recovering the original sources s_m before the mixing took place.

*E-mail: martin.kermit@hiof.no. Corresponding author

†E-mail: oliver.tomic@matforsk.no

2.2. ICA-model of dependent sources

The standard ICA model assumes independent sources, which is not always the case. An alternative ICA-algorithm that deals with dependent sources called dependent component analysis (DCA) has been proposed [6]. Here, a simpler approach is proposed.

Suppose that N observations $x_1, x_2, \dots, x_n, \dots, x_N$ of M dependent sources $d_1, d_2, \dots, d_m, \dots, d_M$ have been linearly mixed together as

$$\mathbf{x} = \mathbf{A}\mathbf{d}, \quad (4)$$

where \mathbf{A} is a full rank $N \times M$ scalar matrix containing the coefficients a_{nm} of the mixing system. The M dependent sources d_1, \dots, d_M of K are again assumed to consist of K independent sources $s_1, s_2, \dots, s_k, \dots, s_K$ that have been linearly mixed together by the sampling system as

$$\mathbf{d} = \mathbf{B}\mathbf{s}, \quad (5)$$

where \mathbf{B} is a full rank $M \times K$ scalar matrix containing the coefficients b_{mk} of the mixing system governed by the nature of the measuring device. The observed signals are thus generated by a double mixture of independent sources,

$$\mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{s}), \quad (6)$$

where \mathbf{A} is a mixture of signals, and \mathbf{B} is the mixture introduced by the sampling system. The goal is to estimate an inverse of \mathbf{B} giving a system \mathbf{V} such that

$$\mathbf{d} = \mathbf{V}\mathbf{s}, \quad (7)$$

thus recovering the original sources s_k before the mixing by the sampling system took place.

Standard ICA-algorithms estimate the sources are under the assumption of a single mixture procedure as described in the standard ICA-model in Eq. (2). In the case of dependent sources, as in Eq. (6), standard ICA-algorithms will produce an unmixing system \mathbf{W} given by

$$\mathbf{W} = (\mathbf{A}\mathbf{B})^{-1}, \quad (8)$$

thus incorporating the sensor dependencies introduced by \mathbf{B} to the unmixing of \mathbf{A} . As a result, the unmixing system \mathbf{W} will not be an accurate estimate of \mathbf{A}^{-1} . This is because the standard ICA-model assumes $\mathbf{B} = \mathbf{I}$, thus no sensor dependencies present. If the choice $\mathbf{A} = \mathbf{I}$ is used, the unmixing system in Eq. (8) becomes

$$\mathbf{W} = (\mathbf{A}\mathbf{B})^{-1} = (\mathbf{I}\mathbf{B})^{-1} = \mathbf{B}^{-1} = \mathbf{V}, \quad (9)$$

and may thus be stored for further reference of the sensor dependencies. When the system again is to measure a process where a mixture of signals is present, that is $\mathbf{A} \neq \mathbf{I}$, an unmixing system \mathbf{W} following Eq. (8) is estimated. The true signal mixture \mathbf{A} can then be calculated using the sensor dependent reference matrix \mathbf{B} by

$$\mathbf{A} = \mathbf{W}^{-1}\mathbf{B}^{-1}. \quad (10)$$

3. DEPENDENCE AND INDEPENDENCE

The extended ICA-model just described and its proposed solution, is only tractable when sensor dependencies are fixed, and do not vary with different input signals. In this section, a set of test signals with known dependencies is introduced and explored. These test signals lead to an alternative way to estimate sensor dependencies when the extended ICA-model does not hold.

3.1. Test signals

Sinusoidal signals are traditionally used as basis functions in frequency analysis. Such a set of basis functions are easy to construct and modify. Due to their property of being mutually orthogonal at specific frequencies only, sinusoidal signals are well suited as test signals since the degree of correlation can be controlled. We define the following,

$$\begin{aligned} \mathbf{x} &= \sin(2\pi\omega_1 t), & t \in T \\ \mathbf{y} &= \sin(2\pi\omega_2 t), & t \in T \end{aligned} \quad (11)$$

where ω_1 and ω_2 denotes the angular frequencies of each of the signals. The correlation coefficient, c , between the two signals \mathbf{x} and \mathbf{y} is given by their normalized inner product,

$$c = \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{y}}{\|\mathbf{y}\|} \right\rangle, \quad (12)$$

where $\|\cdot\|$ denotes the vector norm.

3.2. Correlation between the test signals

From the test signals chosen in Eq. (11), we chose ω_1 to be constant $\omega_1 = 0.4$. The time interval T is set to be equal to two full periods for the sinusoidal signal \mathbf{x} , so that $T = [0, 5]$. For the other test signal, \mathbf{y} , we let ω_2 vary in an interval Ω defined as the range $\Omega = [\omega_1, 2\omega_1]$. The correlation between the two signals \mathbf{x} and \mathbf{y} will thus depend on the chosen value for ω_2 . The coefficient of correlation will have its maximum value, $c = 1$, when ω_2 has its lowest value in the interval Ω , namely $\omega_2 = \omega_1$. Thus, the signals are equal, $\mathbf{x} = \mathbf{y}$. With increasing values for ω_2 , the correlation, c , between the two signals will decrease as the two signals become more independent of each other.

Fig. 1 shows in the top box the correlation coefficient c between \mathbf{x} and \mathbf{y} as a function of the angular frequency ω_2 of the signal \mathbf{y} . It can be seen that the correlation between \mathbf{x} and \mathbf{y} are zero at regularly values for ω_2 , namely

$$\omega_2 = k\omega_1/4, \quad k = 4, 5, \dots \quad (13)$$

At these points, the signals \mathbf{x} and \mathbf{y} are uncorrelated and thus mutually orthogonal to each other. It can be shown that the correlation coefficient, c , as a function of ω_2 follows

$$c(\omega_2) \propto \frac{\sin(\omega_2 - \omega_1)}{(\omega_2 - \omega_1)} - \frac{\sin(\omega_2 + \omega_1)}{(\omega_2 + \omega_1)}. \quad (14)$$

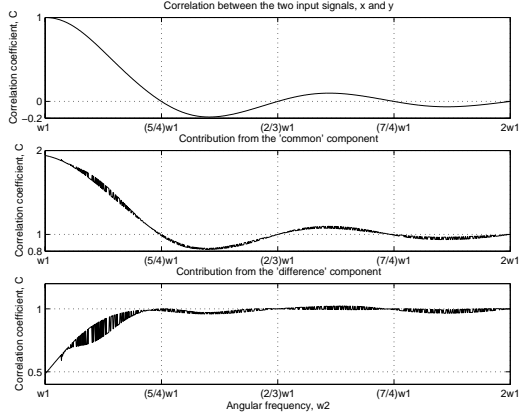


Fig. 1. *Upper:* The correlation coefficient c between the two x and y defined in Eq. (11) as a function of the angular frequency ω_2 of the signal y . *Middle:* The total contribution from the "common" part, c_{common} . *Lower:* The total contribution from the "differ" part, c_{differ} .

3.3. ICA applied to the test signals

As the correlation between the two test signals x and y has been established, we now turn our focus on what to expect when ICA is applied to the test signals¹. It should be evident that when $\omega_2 = k\omega_1/4$, $k = 4, 5, \dots$, x and y are orthogonal and does not share any common source signal. According to the ICA-model, this implies the special case where the observation signals equals the source signals, thus $\mathbf{A} = \mathbf{I}$. The particular situation where $\omega_2 = 2\omega_1$ is indicated in Fig. 2.

If the value of ω_2 is chosen from the lower region of Ω , thus close to the value of ω_1 , high correlation between x and y will result. According to the ICA-model, x and y are considered to be mixtures of some unknown source signals, s_1 and s_2 , and $\mathbf{A} \neq \mathbf{I}$. By the application of an ICA-transform to the highly correlated signals x and y , the resulting source signals s_1 and s_2 will be mutually independent and uncorrelated, and not necessarily mimic the behaviour of any of the input signals x and y .

An example of an ICA-transform of x and y using $\omega_2 = (101/100)\omega_1$ is shown in Fig. 3. As seen from the upper plot in the figure, the two signals are visually similar near $t = 0$, and differs increasingly with increasing t . The middle plot in Fig. 3 shows the first independent component, s_1 extracted from x and y . It is seen that the shape of this component is approximately sinusoidal with a constant amplitude and with an oscillation frequency different from both ω_1 and ω_2 . The second independent component, s_2 plotted in the lower box of Fig. 3, shows us that the increasing difference between the input signals x and y is to a certain

¹All ICA carried out in this text is performed with the FastICA Matlab package provided by Hyvärinen [7, 8]. FastICA was used with the deflation approach and $g(u) = u^3$ with no dimensionality reduction.

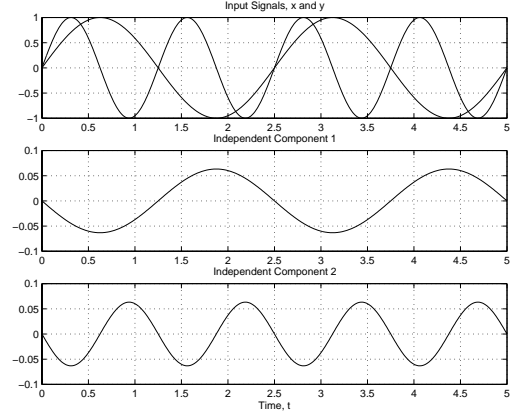


Fig. 2. *Upper:* The two input signals to the ICA-model, x and y , where $\omega_2 = 2\omega_1$, thus indicating that the signals are orthogonal. *Middle:* The first estimated independent component which is approximately equal to a scaled version of x . *Lower:* The second estimated independent component which is approximately equal to a scaled version of y .

degree reflected with the increasing amplitude of s_1 .

The mixing matrix estimated from the ICA-transform in this example suggests that s_1 contributes to roughly 75% of both the two input signals x and y , while the rest is contributed by s_2 . This result suggests that the first component, s_1 , contains a signal part that is "common" in both input signals x and y , while s_2 represent information regarding the "differ" between the signals. The suggestion can be verified by calculating the correlation coefficients between the input signals and the source signals. Thus, the correlation between the source signal s_1 and the input x gives the amount of contribution from the "common" part, while the correlation between s_2 and the input x , the "differ" part yield.

The sum of the correlation coefficients calculated for both input signals between x and y and the source signal s_1 will thus give the total contribution from the "common" part. A similar argument applies for the "difference" part, s_2 . Due to the scaling and permutation property of the ICA-transform, normalization and absolute values are required, so the applied definition is

$$\begin{aligned} c_{\text{common}} &= \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{s}_1}{\|\mathbf{s}_1\|} \right\rangle + \left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|}, \frac{\mathbf{s}_1}{\|\mathbf{s}_1\|} \right\rangle \\ c_{\text{differ}} &= \left\langle \frac{\mathbf{x}}{\|\mathbf{x}\|}, \frac{\mathbf{s}_2}{\|\mathbf{s}_2\|} \right\rangle + \left\langle \frac{\mathbf{y}}{\|\mathbf{y}\|}, \frac{\mathbf{s}_2}{\|\mathbf{s}_2\|} \right\rangle. \end{aligned} \quad (15)$$

Fig. 1 shows plots of c_{common} and c_{differ} for varying $\omega_2 \in \Omega$. The upper box shows the correlation coefficient c between x and y as a function of the angular frequency ω_2 of the signal y . The middle box shows the total contribution from the "common" part, while the lower plot indicates the total contribution from the "differ" part. It is clearly seen that the "common" correlation follows the correlation between x and y . The contribution from the "differ" increases

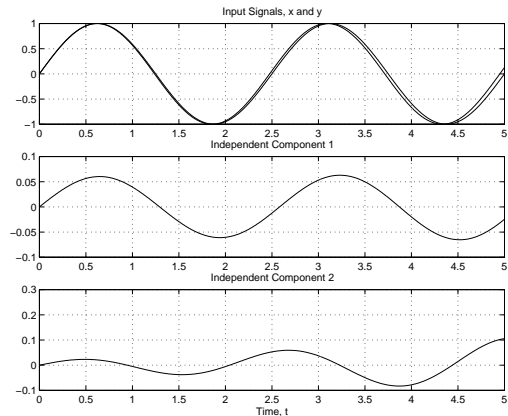


Fig. 3. *Upper:* The two input signals to the ICA-model, \mathbf{x} and \mathbf{y} , where $\omega_2 = (101/100)\omega_1$, thus indicating that the signals are highly correlated. *Middle:* The first independent component, s_1 extracted from \mathbf{x} and \mathbf{y} . It is seen that the shape of this component is approximately sinusoidal with a constant amplitude. *Lower:* The second independent component, s_2 shows that the increasing difference between the input signals \mathbf{x} and \mathbf{y} is to a certain degree reflected with the increasing amplitude of s_1 .

as the correlation between the inputs decreases. It can be noticed that $c_{\text{common}} = 1$ and $c_{\text{differ}} = 1$ for orthogonal inputs, $\omega_2 = k\omega_1/4$, $k = 4, 5, \dots$, which is reasonable since the two source signals then represent each of the inputs, \mathbf{x} and \mathbf{y} .

3.4. ICA applied to mixed test signals

It will now be illustrated how a mixture of highly correlated test signals can be estimated using the procedure described in section 2.2. The test signals used are \mathbf{x} and \mathbf{y} where $\omega_2 = (101/100)\omega_1$ as illustrated before in Fig. 3. Now, a mixture \mathbf{A} is introduced between the signals given by

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}. \quad (16)$$

The two mixed signals are shown in the top box of Fig. 4. In the two boxes below are plotted the result of the ICA-transform. As seen, the estimated source signals are approximately the same as in the case where no mixing took place (Fig. 3), since the ICA-transform has estimated an unmixing system incorporating both the applied signal mixture \mathbf{A} and the correlation between the signals \mathbf{B} . The correlation matrix \mathbf{B} can be established using the unmixed signals. Thus, Eq. (10) can be used to estimate the signal mixture, which in this case was estimated to be

$$\hat{\mathbf{A}} = \begin{bmatrix} 1.00 & 2.00 \\ 2.99 & 4.01 \end{bmatrix}. \quad (17)$$

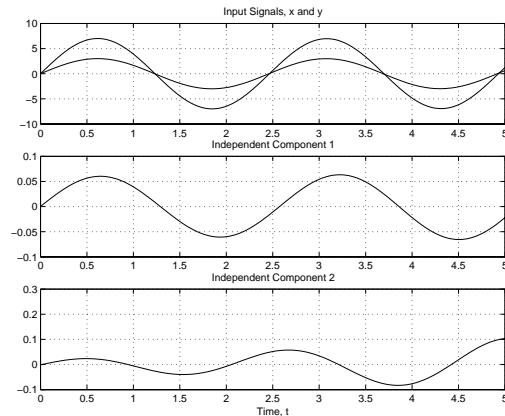


Fig. 4. *Upper:* The test signals mixed according to the mixing matrix in Eq. (16). *Middle:* First independent component, representing the "common" signal part. *Lower:* Second independent component, representing the "differ" part of the signals.

4. APPLICATION TO ODOUR SIGNALS

In this section, a suggestion of how the described ICA model can be used to separate odour signals generated from alcohol solutions and recorded with an electronic nose is presented, as well as giving estimates of the mixing applied to the signals. This process can be viewed as a cocktail party of odours instead of sound signals.

4.1. Electronic nose

With the term electronic nose is understood an array of chemical gas sensors with a broad and partly overlapping sensitivity for measurements of volatile compounds combined with multivariate statistical data processing tools. The electronic nose belong to the category of rapid analysis instruments, allowing non-destructive analysis of vapours at a high rate with sufficient reproducibility and accuracy.

Sensor data acquisition is accomplished by dynamic sampling of head-space gas, i.e. gas from the vial, saturated with volatiles from the sample, pumped over the sensor array. While exposed to the analyte gas, the semi-conducting sensors generate an electronic output signal that is used to determine the sensor response value for the measured sample.

In the work presented in this text, a hybrid gas sensor array manufactured by AppliedSensor Technologies in Linköping, Sweden was used to perform the measurements. The gas sensor array consisted of 22 gas sensors. Each sensor possessed a unique sensitivity profile with varying sensitivity and selectivity towards certain volatile compounds.

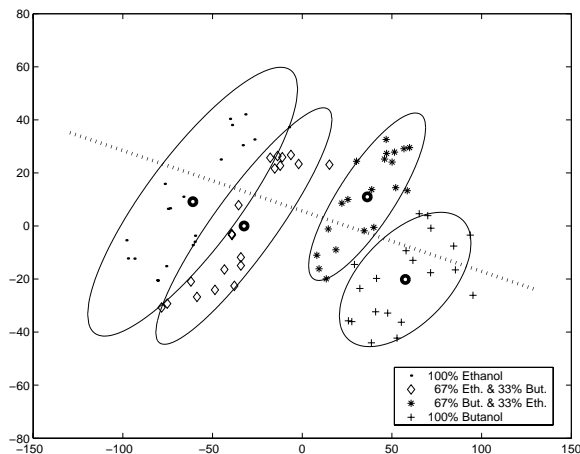


Fig. 5. Score plot showing the linear relationship of the graded concentration between the four solutions: Pure ethanol 0.5 % and butanol 1%, and two mixtures of the solutions with a ratio 1:2 and 2:1.

4.2. Materials and methods

The odours subject to inspection by the proposed scheme should satisfy the two criteria set by the ICA-model. First, the odour signals should be as independent as possible before sampling by the electronic nose. Second, a mixture of two odour components should follow a linear relationship. The last condition is not met if the mixing of gases involves a chemical reaction, thus generating new molecular structures in the gas.

For the following investigation, four different solutions with butanol 1% and ethanol 0.5% were used, since alcohol solutions meet the mentioned criteria. The concentration level of the solutions were chosen to fit in the linear response area of the electronic nose. Two of the solutions consisted of only ethanol and butanol, and the other two were mixtures using a ratio of 1:2 and 2:1 for both ethanol and butanol. The first two solutions will in the rest of this text be referred to as the pure alcohol solutions, while the latter are called mixtures. By this procedure, the four solutions then represented a graded decrease in concentration for both ethanol and butanol. This graded concentration decrease was verified using the standard data analysis tool for electronic noses, principal component analysis (PCA). Each of the four solutions were measured 20 times for each of the 22 sensors in the array. Selected features from the response curves of all the sensors were used to create a PCA score-plot from the response signals of the electronic nose. The score-plot, using the two most prominent principal components, is shown in Fig. 5. It is seen that the four solutions form clusters that to some extent are separable to each other. Also indicated is a line fitted according to the cluster means showing the linear relationship between the four solutions,

thus verifying that the responses are in the linear response area of the electronic nose.

For the rest of this work, only 2 of the 22 sensors in the array were picked for the further analysis according to their level of response to the alcohol solutions. Both sensors were of the MOSFET² transistor type. Typical response signals for the two sensors (numbered 12 and 13) is shown in Fig. 6. From the traces shown, it should be clear that the sensor responses are highly correlated for all solutions.

4.3. Separation of odours using ICA

The sensor responses from the two sensors to the four alcohol solutions described previously were ICA-transformed one at the time, thus giving two independent components for each of the four solutions. The estimated independent components are shown in Fig. 7. It can be seen from the traces that for each of the solutions one of the components indicate strong relationship with the sensor response (Fig. 6) of the major ingredient in the measured solution as opposed to the other component. As described for the test signals in Sec. 3.3, this result indicates a "common" and a "differ" component.

The goal of this investigation is to establish the ratio with which the two mixed alcohol solutions were mixed, using only the available source signals estimated from the pure alcohol solutions by the ICA-transform. Traditionally, the calculation of such mixing ratios of odours has been a challenging problem to the electronic nose community. This has been due to both response variability of the sampling instruments and the lack of suitable analysis tools.

Due to the correlation that clearly exists between the sensor responses, the unmixing system estimated from the ICA-transform will incorporate both the sensor dependencies, \mathbf{B} , and the applied mixtures of the solutions, \mathbf{A} . Since the pure solutions of both ethanol 0.5% and butanol 1.0% are available, the dependency matrix \mathbf{B} can be estimated according to the scheme outlined in section 2.2. Unfortunately, the sensor dependencies differs according to the odour it is exposed to and can not be linearly described. The estimated dependency matrix for the ethanol solution, \mathbf{B}_e is not equal to the estimated dependency matrix using the ethanol solution, \mathbf{B}_b . If instead, a dependency matrix using an average of \mathbf{B}_e and \mathbf{B}_b where their columns are normalized is used, a suitable dependency matrix can be formed.

A more direct approach in estimating the contribution from the two mixed alcohol solutions is the calculation of c_{common} and c_{differ} as defined in Eq. (15). These are calculated for both of the pure alcohol solutions and added together, giving the total contribution from each of the alcohols.

²MOSFET (metal-oxide-semiconductor field-effect transistor) gas sensors consist of three layers; a doped silicon semiconductor, a thick oxide layer (SiO₂) as insulator and on top a thin catalytic metal layer.

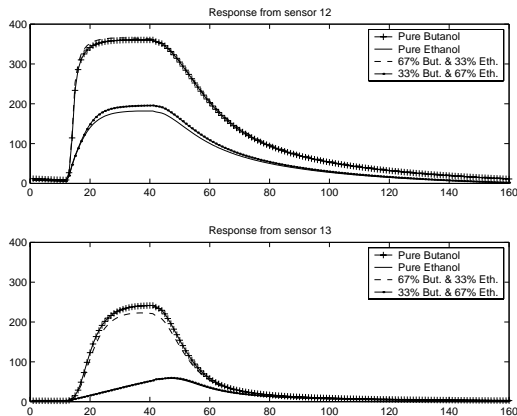


Fig. 6. Typical response signals as functions of time for two selected sensors when exposed to the four alcohol solutions.

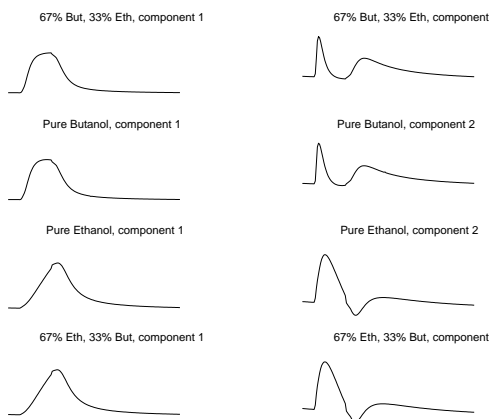


Fig. 7. Source signals from the four alcohol solutions estimated by the ICA-transform. "Common" components are shown left, while "differ" components are shown right.

Table 1 shows the estimated mixing ratios averaged over 11 odour signals for each solution for the two mixed alcohol solutions by both the dependency matrix approach and the direct method with correlation calculation. It can be seen that the averaged estimates are quite close to the true ratios, 1:2 and 2:1 for both approaches, while the variance is still considerably large. This variance may be explained by the large variability in the responses produced by electronic noses.

5. CONCLUSIONS

ICA has been applied as a novel tool in the analysis of data produced by odours sampled by an electronic nose. It has been shown that sensor dependencies present in the measurement process may be overcome by the use of ICA to either construct a dependency matrix in the linear dependency

| Est. content | Eth.-But. ratio 2:1 | | Eth.-But. ratio 1:2 | |
|---------------|---------------------|----------|---------------------|----------|
| | Eth. (%) | But. (%) | Eth. (%) | But. (%) |
| Dep. matrix | 63 ± 13 | 37 ± 13 | 36 ± 14 | 64 ± 14 |
| Cor. approach | 66 ± 13 | 34 ± 13 | 36 ± 12 | 64 ± 12 |

Table 1. Classification results of the two mixed alcohol solutions using dependency matrix and the correlation approach. Indicated are percentage of the estimated content.

case, or by a direct estimate of the correlation between the observed signals and some reference signals recorded without mixing. The described methods has successfully been applied to a real world application, estimating the mixing ratio with which two alcohol solutions were mixed, though with large variance. In the future, the large variance in the results should be decreased, and other types of odours involving more than two compounds should be tested. Further investigation using ICA on odour signals should follow.

6. REFERENCES

- [1] Te-Won Lee, *Independent Component Analysis - Theory and applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1998.
- [2] A. J. Bell and Terrence J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, pp. 1129–1159, 1995.
- [3] P. Comon, "Independent component analysis - a new concept?," *Signal Processing*, vol. 36, pp. 287–314, 1994.
- [4] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural Networks*, vol. 13, no. 4-5, pp. 411–430, 2000.
- [5] C. Jutten and J. Herault, "Blind separation of sources, part i: An adaptive algorithm based on neuromimetic architecture," *Signal Processing*, vol. 24, pp. 1–10, 1991.
- [6] Allan Kardec Barros, *The Independence Assumption: Dependent Component Analysis*, chapter 4, pp. 63–70, Perspectives in Neural Computing. Springer, 2000, Advances in Independent Component Analysis.
- [7] A. Hyvärinen and E. Oja, "A fast fixed-point algorithm for independent component analysis," *Neural Computation*, vol. 9, no. 7, pp. 1483–1492, 1997.
- [8] A. Hyvärinen, "The fixed-point algorithm and maximum likelihood estimation for independent component analysis," *Neural Processing Letters*, vol. 10, no. 1, pp. 1–5, 1999.